

WEAK CONVERGENCE OF CD KERNELS: A NEW APPROACH ON THE CIRCLE AND REAL LINE

Brian Simanek¹

ABSTRACT. Given a probability measure μ supported on some compact set $K \subseteq \mathbb{C}$ and with orthonormal polynomials $\{p_n(z)\}_{n \in \mathbb{N}}$, define the measures

$$d\mu_n(z) = \frac{1}{n+1} \sum_{j=0}^n |p_j(z)|^2 d\mu(z)$$

and let ν_n be the normalized zero counting measure for the polynomial p_n . If μ is supported on a compact subset of the real line or on the unit circle, we provide a new proof of a 2009 theorem due to Simon that for any fixed $k \in \mathbb{N}$ the k^{th} moment of ν_{n+1} and μ_n differ by at most $O(n^{-1})$ as $n \rightarrow \infty$.

Keywords: Orthogonal polynomials, Reproducing Kernel, Balayage

Mathematics Subject Classification: 42C05, 60B10

1. INTRODUCTION

Given a probability measure μ on \mathbb{C} with infinite and compact support, we can form the sequence $\{p_n(z)\}_{n \in \mathbb{N}}$ of orthonormal polynomials satisfying

$$\int_{\mathbb{C}} \overline{p_n(z)} p_m(z) d\mu(z) = \delta_{n,m}$$

and normalized so that each p_n has positive leading coefficient κ_n . With this sequence, we define

$$K_n(z, \zeta; \mu) = \sum_{j=0}^n \overline{p_j(\zeta)} p_j(z),$$

the so-called *Reproducing Kernel* for polynomials of degree n . We assign it this name because of the reproducing property, namely that if Q is any polynomial of degree at most n then

$$Q(w) = \int_{\mathbb{C}} Q(z) K_n(w, z; \mu) d\mu(z).$$

With this notation, we can define the probability measures

$$d\mu_n = \frac{K_n(z, z; \mu)}{n+1} d\mu$$

¹Mathematics MC 253-37, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: bsimanek@caltech.edu. Supported in part by an NSF GRFP grant.

for each $n \in \mathbb{N}$.

If we write $p_n(z) = \kappa_n \prod_{j=1}^n (z - z_j^{(n)})$ (the $z_j^{(n)}$ need not be distinct), then we define the measures

$$d\nu_n = \frac{1}{n} \sum_{j=1}^n \delta_{z_j^{(n)}}$$

for each $n \in \mathbb{N}$. In [3], Simon proved the following theorem:

Theorem 1.1. [3] *Let $N(\mu) = \sup\{|z| : z \in \text{supp}(\mu)\}$. For any $k \in \mathbb{N}$, we have*

$$(1.1) \quad \left| \int_{\mathbb{C}} z^k d\mu_n(z) - \int_{\mathbb{C}} z^k d\nu_{n+1}(z) \right| \leq \frac{2kN(\mu)^k}{n+1}.$$

From this theorem, Simon deduces the following important corollary. Suppose C is a circle centered at 0 with radius larger than $N(\mu)$. Let $\hat{\mu}_n$ denote the balayage (see Theorem II.4.1 in [2]) of the measure μ_n onto C and similarly define $\hat{\nu}_n$. It follows from Theorem 1.1 that

$$w\text{-}\lim_{\substack{n \rightarrow \infty \\ n \in \mathcal{N}}} \hat{\nu}_{n+1} = \sigma \quad \Longleftrightarrow \quad w\text{-}\lim_{\substack{n \rightarrow \infty \\ n \in \mathcal{N}}} \hat{\mu}_n = \sigma,$$

where σ is a probability measure on C and $\mathcal{N} \subseteq \mathbb{N}$ is a subsequence.

The proof of the above theorem in [3] relies on the relationship between the polynomials $\{p_n\}_{n \in \mathbb{N}}$ and the eigenvalues of the operator M_z acting on $L^2(\mu)$ by $M_z(f(z)) = zf(z)$. In this paper, we will provide a new proof of this theorem when $\text{supp}(\mu) \subseteq \mathbb{R}$ or $\text{supp}(\mu) \subseteq \partial\mathbb{D} := \{z : |z| = 1\}$. The key idea will be to look at Prüfer phases of the appropriate ratio of the orthonormal polynomials.

If μ is supported on $\partial\mathbb{D}$, we define $\eta_n(\theta) : [0, 2\pi] \rightarrow \mathbb{R}$ to be a continuous function so that

$$(1.2) \quad e^{i\eta_n(\theta)} = \frac{p_{n+1}(e^{i\theta})}{p_{n+1}^*(e^{i\theta})},$$

where $p_{n+1}^*(z) = z^{n+1} \overline{p_{n+1}(\bar{z}^{-1})}$ (so that the right hand side of (1.2) is a Blaschke product). If μ is supported on \mathbb{R} (we always assume compact support), then we may define $\theta_n(x) : \mathbb{R} \rightarrow (-\pi/2, \infty)$ to be a continuous function so that

$$(1.3) \quad \tan(\theta_n(x)) = \frac{a_n p_n(x)}{p_{n-1}(x)},$$

(see Proposition 6.1 in [1]) where a_n is a positive real number so that p_{n-1} and $a_n p_n$ have the same leading coefficient. In our proofs, we will use the functions η_n and θ_n (more precisely their derivatives) to obtain measures that approximate the measure μ in a sense suitable for our purposes.

More precisely, two approximating measures will enter. In the unit circle case, we define

$$(1.4) \quad d\mu^n = |p_{n+1}(e^{i\theta})|^{-2} \frac{d\theta}{2\pi}$$

for each $n \in \mathbb{N}$. The measure μ^n (called the n^{th} Bernstein-Szegő measure) is in fact a probability measure on $[0, 2\pi]$ and it induces a measure on $\partial\mathbb{D}$ with the same first n moments

- and hence the same first n orthonormal polynomials - as μ (this follows from Theorems 1.7.8 and 1.5.5 in [6]). In the real line case, we define

$$(1.5) \quad d\rho_n = \frac{dx}{\pi(a_{n+1}^2 p_{n+1}(x)^2 + p_n(x)^2)}$$

as in Theorem 2.1 in [4]. It follows from equation (2.7) in [4] that $d\rho_n$ is a probability measure and

$$(1.6) \quad \int_{\mathbb{R}} x^\ell d\rho_n(x) = \int_{\mathbb{R}} x^\ell d\mu(x) \quad , \quad \ell = 0, 1, \dots, 2n.$$

In the next section we provide our new proof of Theorem 1.1 when μ is supported on the unit circle. In Section 3 we consider μ supported on the real line and prove Theorem 1.1 with the right hand side of (1.1) replaced by $O(n^{-1})$.

2. THE UNIT CIRCLE CASE

Our goal in this section is to provide a new proof of Theorem 1.1 when μ is supported on the unit circle. We begin our proof by noting that the theorem is equivalent to the statement that the moments of the signed measures $d\hat{\nu}_{n+1} - d\mu_n$ converge 0 at a certain rate where $\hat{\nu}_n$ is the balayage of the measure ν_n onto $\partial\mathbb{D}$. It is easy to check that (see equation (8.2.8) in [6])

$$d\hat{\nu}_{n+1} = \frac{1}{n+1} \sum_{j=1}^{n+1} \frac{1 - |z_j^{(n+1)}|^2}{|e^{i\theta} - z_j^{(n+1)}|^2} \frac{d\theta}{2\pi}.$$

If we define $\eta_n : [0, 2\pi] \rightarrow \mathbb{R}$ as in (1.2) above, then equation (6.10) in [7] implies that

$$\frac{d}{d\theta} \eta_n(\theta) = \sum_{j=1}^{n+1} \frac{1 - |z_j^{(n+1)}|^2}{|e^{i\theta} - z_j^{(n+1)}|^2}.$$

Furthermore, equation (10.8) in [1] tells us that

$$\frac{d}{d\theta} \eta_n(\theta) = \frac{K_n(e^{i\theta}, e^{i\theta}; \mu)}{|p_{n+1}(e^{i\theta})|^2}$$

so we conclude that

$$d\hat{\nu}_{n+1} = \frac{K_n(e^{i\theta}, e^{i\theta}; \mu)}{n+1} d\mu^n(\theta).$$

Therefore, if $k \in \mathbb{N}$, we can write

$$\int_{\mathbb{D}} z^k d\nu_{n+1}(z) - \int_{\partial\mathbb{D}} z^k d\mu_n(z) = \frac{1}{n+1} \sum_{j=0}^n [\langle p_j(z), z^k p_j(z) \rangle_{\mu^n} - \langle p_j(z), z^k p_j(z) \rangle_{\mu}].$$

Since μ and μ^n have the same first n moments, at most k of these summands are non-zero and each non-zero summand has absolute value at most 2. We have therefore proven

$$\left| \int_{\mathbb{D}} z^k d\mu_n(z) - \int_{\partial\mathbb{D}} z^k d\nu_{n+1}(z) \right| \leq \frac{2k}{n+1}$$

exactly as in Theorem 1.1.

Example. Let μ be the normalized arclength measure on the unit circle. In this case we have $p_n(z) = z^n$ for all n and $\mu_n = \mu$ for all n . The measures ν_n are all simply the point mass at 0 with weight 1. This example illustrates the fact that in general, the measures μ_n and ν_n need not resemble each other as measures on $\overline{\mathbb{D}}$, so it really is important that we consider the balayage.

3. THE REAL LINE CASE

Our goal in this section is to provide a new proof of Theorem 1.1 when μ is supported on a compact subset of the real line and with the right hand side of (1.1) replaced by $O(n^{-1})$ where the implied constant depends on k . There is a proof of this result due to Totik, also appearing in [3], but with the right hand side of (1.1) replaced by $o(1)$ (though it can be modified to produce the same $O(n^{-1})$ discrepancy estimate for the moments as in (1.1) above). Totik's proof uses Gaussian quadratures and the monotonicity (in n) of the sequence $K_n(x, x; \mu)$ to establish the weak convergence result for all polynomials that are positive on the convex hull of the support of μ . The proof we present here will be analogous to the proof in Section 2 and will rely on the sequence of approximating measures ρ_n (see (1.5) above). We will make use of formula (3.1) below, which relates a set of perturbed zero-counting measures to a set of perturbed quadrature measures. Combining this with an interlacing property will allow us to derive the $O(n^{-1})$ estimate in (1.1).

Our computation will be a bit longer than in the unit circle case partly because in Section 2, the most difficult calculation was already done for us in [1] and partly because the high moments of the measure ρ_n defined in equation (1.5) are infinite, so we need to use a cutoff function.

Let us assume μ has support contained in $[-M, M]$ and define

$$\tau(x) = \chi_{[-M-1, M+1]}(x).$$

Corresponding to μ there is a Jacobi matrix J , which is the matrix of multiplication by x in the Hilbert space $L^2(\mu)$ with respect to the basis given by the orthonormal polynomials. For any $\lambda \in \mathbb{R}$, we will let $\mu_{n,\lambda}$ be the spectral measure corresponding to the Jacobi matrix $J_n + \lambda \langle e_n, \cdot \rangle e_n$ and the vector e_1 where J_n is the upper left $n \times n$ block of J (see Section 6 in [5]). Notice that $\mu_{n,\lambda}$ is supported on n distinct points and by Corollary 6.3 in [5], the points in the support of $\mu_{n,\lambda}$ interlace for distinct values of λ . Let $\nu_{n,\lambda}$ be the measure placing weight n^{-1} on each point in the support of $\mu_{n,\lambda}$ (so that $\nu_{n,0} = \nu_n$). It follows from formula (6.16) in [5] that

$$(3.1) \quad \frac{1}{n} d\mu_{n,\lambda} = \frac{1}{K_{n-1}(x, x; \mu)} d\nu_{n,\lambda}.$$

Therefore for any fixed $k \in \mathbb{N}$, we have

$$(3.2) \quad \int_{\mathbb{R}} x^k \tau(x) d\nu_{n+1,\lambda} = \frac{1}{n+1} \int_{\mathbb{R}} x^k \tau(x) K_n(x, x; \mu) d\mu_{n+1,\lambda}.$$

After taking a suitable average (in λ), the expression on the left-hand side of (3.2) approximates the k^{th} moment of ν_{n+1} as $n \rightarrow \infty$ while the right-hand side approximates the k^{th} moment of μ_n as $n \rightarrow \infty$. Indeed, our first step is to integrate the left hand side of (3.2) from $-\infty$ to ∞ with respect to $\frac{d\lambda}{\pi(1+\lambda^2)}$. Notice that for any value of λ , at most one point in the support of $\nu_{n+1,\lambda}$ lies outside $[-M-1, M+1]$ because of the interlacing property. Therefore, we have

$$(3.3) \quad \int_{-\infty}^{\infty} \int_{\mathbb{R}} x^k \tau(x) d\nu_{n+1,\lambda}(x) \frac{d\lambda}{\pi(1+\lambda^2)} = \int_{\mathbb{R}} x^k \tau(x) d\nu_{n+1,0}(x) + O(n^{-1})$$

as $n \rightarrow \infty$.

If we integrate the right hand side of (3.2) in the same way, this becomes

$$(3.4) \quad \frac{1}{n+1} \int_{\mathbb{R}} x^k \tau(x) K_n(x, x; \mu) d\rho_n(x)$$

by Theorem 2.1 in [4]. Notice that this integral would be infinite without the cut-off function τ . As an aside, we note that by Proposition 6.1 in [1], (3.4) is just

$$\frac{1}{n+1} \int_{\mathbb{R}} x^k \tau(x) \frac{1}{\pi} \frac{d\theta_{n+1}(x)}{dx} dx,$$

which is why we call this the analog of the proof in Section 2. Notice that for any fixed $m \leq n$ we have

$$\int_{\mathbb{R}} x^k \tau(x) |p_m(x)|^2 d\rho_n(x) \leq (M+1)^k.$$

This follows from the fact that p_m is also the degree m orthonormal polynomial for the measure $d\rho_n$ by (1.6). Therefore, we can rewrite (3.4) as

$$\frac{1}{n+1} \int_{\mathbb{R}} x^k \tau(x) K_{n-k}(x, x; \mu) d\rho_n(x) + O(n^{-1})$$

as $n \rightarrow \infty$. We can rewrite this again as

$$(3.5) \quad \frac{1}{n+1} \int_{\mathbb{R}} x^k K_{n-k}(x, x; \mu) d\rho_n(x) - \frac{1}{n+1} \int_{|x|>M+1} x^k K_{n-k}(x, x; \mu) d\rho_n(x) + O(n^{-1})$$

as $n \rightarrow \infty$. Notice that $x^k K_{n-k}(x, x; \mu)$ is a polynomial of degree $2n-k$ while the denominator of the weight defining the measure ρ_n is a polynomial of degree $2n+2$. Therefore, both integrals in (3.5) are finite. The first term in (3.5) is equal to

$$\frac{1}{n+1} \int_{\mathbb{R}} x^k K_n(x, x; \mu) d\mu(x) + O(n^{-1}) = \int_{\mathbb{R}} x^k d\mu_n(x) + O(n^{-1})$$

as $n \rightarrow \infty$ again by (1.6). We will be finished if we can show that the second term in (3.5) tends to 0 like $O(n^{-1})$ as $n \rightarrow \infty$ and for this it suffices to put a uniform bound on

$$(3.6) \quad \int_{|x|>M+1} x^k K_{n-k}(x, x; \mu) d\rho_n(x).$$

To do this, we rewrite (3.6) as

$$\int_{-\infty}^{\infty} \int_{|x|>M+1} x^k K_{n-k}(x, x; \mu) d\mu_{n+1,\lambda}(x) \frac{d\lambda}{\pi(1+\lambda^2)}.$$

Recall that for each fixed λ , at most one point in the support of $\mu_{n+1,\lambda}$ has absolute value larger than $M+1$. Let us denote this point (if it exists) by $x_{n+1,\lambda}$. Therefore, the above integral is just

$$(3.7) \quad \int_A x_{n+1,\lambda}^k \frac{K_{n-k}(x_{n+1,\lambda}, x_{n+1,\lambda}; \mu)}{K_n(x_{n+1,\lambda}, x_{n+1,\lambda}; \mu)} \frac{d\lambda}{\pi(1+\lambda^2)},$$

where we used (3.1) and the integral is taken over some set $A \subseteq \mathbb{R}$ such that $x_{n+1,\lambda}$ exists if and only if $\lambda \in A$. Using the Christoffel Variational Principle (Theorem 9.2 in [5]), it is easily seen that

$$\frac{K_{n-k}(x_{n+1,\lambda}, x_{n+1,\lambda}; \mu)}{K_n(x_{n+1,\lambda}, x_{n+1,\lambda}; \mu)} \leq \left(\frac{M}{|x_{n+1,\lambda}|} \right)^{2k}.$$

Therefore, we can bound (3.7) from above in absolute value by

$$\int_A \frac{M^{2k}}{|x_{n+1,\lambda}|^k} \frac{d\lambda}{\pi(1+\lambda^2)},$$

which is uniform in n since $|x_{n+1,\lambda}| > M+1$. This completes the proof.

Acknowledgments. It is a pleasure to thank my advisor Barry Simon for many useful comments and suggestions.

REFERENCES

- [1] Y. Last, B. Simon, *Fine structure of the zeros of orthogonal polynomials IV: a priori bounds and clock behavior*, Comm. Pure Appl. Math. 61 (2008) 486-538.
- [2] E. B. Saff and V. Totik *Logarithmic Potentials with External Fields*, Grundlehren der Mathematischen Wissenschaften, Band 316, Springer, Berlin-Heidelberg, 1997.
- [3] B. Simon, *Weak convergence of CD kernels and applications*, Duke Math. J. 146 (2009) 305-330.
- [4] B. Simon, *Orthogonal polynomials with exponentially decaying recursion coefficients*, Probability and Mathematical Physics (eds. D. Dawson, V. Jaksic, and B. Vainberg), CRM Proc. and Lecture Notes, 42 (2007), 453-463.
- [5] B. Simon, *The Christoffel-Darboux Kernel*, in "Perspectives in PDE, Harmonic Analysis and Applications," a volume in honor of V. G. Maz'ya's 70th birthday, Proceedings of Symposia in Pure Mathematics, 79 (2008), 295-335.
- [6] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part One: Classical Theory*, American Mathematical Society, Providence, RI, 2005.
- [7] M. Stoiciu, *The statistical distribution of the zeroes of random paraorthogonal polynomials on the unit circle*, J. Approx. Theory 139 (2006), 29-64.